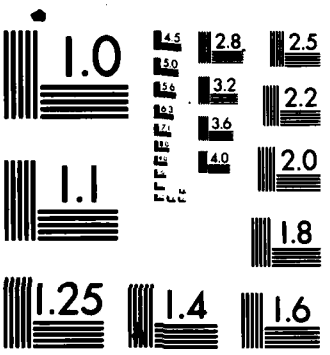
AD-A160 941 AN INFORMATION VALUE METRIC FOR COST DATA STRUCTURES 1/1
(U) OFFICE OF THE SECRETARY OF DEFENSE WASHINGTON DC
DIRECTOR PROGRAM ANALYSIS AND EVALUATION B C JACK 1985
UNCLASSIFIED F/G 5/3 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A160 941

An Information Value Metric For Cost Data Structures

by

Bryan C. Jack
Cost Analyst
OSD(PA&E) Cost Analysis Division
Pentagon, Washington D.C. 20301

Presented at the
19th Annual Department of Defense
Cost Analysis Symposium
Xerox Training Center
Leesburg, Virginia
September 17 - 20 1985

DTIC FILE COPY

DTIC
ELECTE
NOV 6 1985
A

This document has been approved for public release and sale; its distribution is unlimited.

85 11 06 037

An Information Value Metric for Cost Data Structures

Bryan C. Jack
OSD(PA&E)

Efforts in OSD(PA&E) to obtain more detailed total-program-cost estimates for a major weapon system (Program "A") motivated the development of a way to calculate the amount of detail in a cost report. The resulting measure, called the Equivalent Equal-Value Category (EEVC) measure, is derived from a basic formula of the statistical theory of communication. The EEVC measure is applied to cost estimates of Program "A" and of analogous programs, and to pertinent cost reporting outlines. Limitations of the EEVC measure are discussed, its current implementation is described, and areas for its further development are indicated.

Motivation for Developing an Information Value Measure.

OSD reviews cost estimates and other aspects of major DoD weapon systems for processes such as the annual PPBS cycle and DSARC Milestones. When a major program approaches Milestone II, OSD requests that the sponsoring Service provide total life-cycle cost estimates, including RDT&E, procurement, and O&S costs, in sufficient detail for OSD to check the reasonableness of the estimates.

When a certain major weapon system, "Program A," came up for Milestone II, eight categories of RDT&E cost were offered by the "System A" program office; one of these categories contained 55% of the total RDT&E cost. One procurement category, containing 100% of the cost, was offered to OSD. OSD cost analysts felt that the Service could have provided a more detailed cost estimate.

OSD might have "expected" to obtain about as much cost detail about "Program A" as in established standards for weapon system cost reporting, or in cost estimates from programs similar to "Program A." The MIL-STD-881 appendix pertaining to "Program A" contains six Level 2 and 28 Level 3 cost categories. The existing CCDR plan for a similar system, "Program B," contained 16 RDT&E and 27 procurement cost categories. Program cost estimates were on hand for two similar systems: "Program B," (nine RDT&E categories, the largest with 62% of total RDT&E; and five procurement categories, the largest with 46% of total procurement), and "Program C" (19 RDT&E categories, the largest containing 17% of total RDT&E; and 18 procurement categories, the largest with 14% of total procurement cost). Although these examples were not uniform in the number of their cost categories, they generally had more detail than in the original "Program A" cost report.

OSD decided to request more detailed "Program A" cost estimates from the sponsoring Service. No one method was used to develop the categories specified in the OSD requests. In part, the requests were patterned after MIL-STD-881 and similar systems such as programs "B" and "C." The OSD requests were also influenced by the desire to "break up" the aggregated cost categories of the first "Program A" RDT&E cost estimate.

There also was a desire to track progress in this effort to obtain greater cost detail -- to be able to state (in a simpler manner than by reference to a cost outline) how much cost information was on hand, and what was the difference in the amount of cost information between two reports or reporting standards. This desire to express and compare detail in cost reports led to the development of a quantitative measure.

The Equivalent Equal-Value Category (EEVC) Measure of Cost Report Detail.

The resulting measure of cost report detail reflects two features that OSD sought in cost reports it requested from Services: numerous cost categories, and absence of large, highly aggregated cost categories. In the extreme, the "ideal" cost report would have many categories of the same size (or percentage of total cost). The measure of cost report detail should then express the equivalent number of equal-value cost categories. Thus, the measure was called the Equivalent Equal-Value Category (or EEVC) measure.

The EEVC measure of detail in a cost report was derived from a basic formula of the statistical theory of communication. The interdisciplinary link was established by analogy between the cost report's categories, and the set of mutually exclusive possible messages that comprise a communication. (Although the analogy does not lead far into communication theory, it is useful to note that the statistical theory of communication is a branch of applied mathematics, intermediate between stochastic theory (part of abstract probability theory) and communications engineering. [Machol, ed., Systems Engineering Handbook, 1965, McGraw-Hill])

ANALOGY BETWEEN COST REPORTING AND INFORMATION THEORY

COST REPORT	STOCHASTIC MESSAGE
TOTAL COST $C_T = \Sigma_{ci}$	PROBABILITIES OF POSSIBLE MESSAGES $= 1.0 = \Sigma_{pi}$

Three criteria form the basis for the measure of detail in cost reports:

1) The measure should prefer equal-sized cost categories. Assume there is a set of cost reports, each with N categories. The report with the greatest detail is the one whose categories are equal-sized.

2) Disjoint cost categories should be independent of one another. The amount of information, or detail, contributed by one cost category (or group of categories) is independent of how other, disjoint, categories are or are not subdivided.

3) Zero-cost categories should be irrelevant. Although detail may be increased by dividing existing category into two or more non-zero categories, creating categories with zero contents adds nothing to the degree of detail in the cost report.

From these criteria, the following measure of detail in a cost report can be derived:

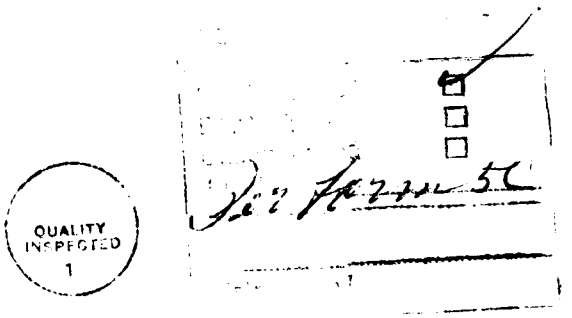
INFORMATION MEASURE FOR COST REPORTS

ASSUME TOTAL COST, $C_T = \Sigma_{ci}$

INFORMATION IN BINARY BITS, $I = -\Sigma \frac{c_i}{C_T} \log_2 (\frac{c_i}{C_T})$

EQUIVALENT EQUAL VALUE CATEGORIES (EEVC) $= 2^I$

The EEVC measure may directly be applied in cost reports where costs are on hand (as opposed to cost outlines, which are generic or do not yet have cost data). The following example comes from the original Service report on RDT&E costs of "Program A." Although the report outline had eight cost categories (and potentially 3 binary "bits" of detail), the inequality of size between the categories meant that there were only 2.17 "bits" of detail, and the equivalent of 4.5 equal-size cost categories. In the EEVC sense, it would have been possible to create a more "informative" report, with only six or seven categories, if some of the smaller categories (such as "launcher", "training", or "system engineering and coordination") were aggregated, and the large "missile" cost category were disaggregated.



A-1

EEVC EXAMPLE: PROGRAM "A" COST-REPORT — RDT&E COSTS

COST CATEGORY	PERCENT OF TOTAL RDT&E	AMOUNT OF DETAIL
Launcher	4.3%	8 Categories
Fire Control	5.3%	
Guidance	14.6%	2.17 "Bits" of Detail
Navigation	8.7%	
Test Instrumentation	5.1%	
Missile	54.8%	4.5 Equivalent Equal-Value Categories (EEVCs)
Training	4.3%	
System Engineering/ Coordination	2.9%	
TOTAL	100%	

The EEVC measure can be extended to cost report outlines that contain no cost data. One way to do this is to assume that each stem of part of the outline contains an equal fraction of the cost in its immediate root. This is called allocating costs per stirpes to the cost outline categories. (An alternative approach would be to assume that each undivided cost category in the outline contains an equal fraction of the total cost. This would be allocating costs per capita to the categories. The per capita approach is not consistent with the EEVC principle of independence of disjoint cost categories, and thus is not used here.) The following example applies the per stirpes principle in calculating the detail imputed from a cost outline:

INFORMATION IMPUTED FROM A BLANK COST OUTLINE

WBS LEVEL 1	WBS LEVEL 2	WBS LEVEL 3	RELATIVE WEIGHT	"BITS" OF DETAIL	EEVCs
TOTAL PROGRAM					
	SUBSYSTEM "A"		(1/2)		
		ITEM 1	1/6	0.431	
		ITEM 2	1/6	0.431	
		ITEM 3	1/6	0.431	
	SUBSYSTEM "B"		(1/2)		
		ITEM 1	1/8	0.375	
		ITEM 2	1/8	0.375	
		ITEM 3	1/8	0.375	
		ITEM 4	1/8	0.375	
			1.0	2.792	6.92

Application of EEVC Measure.

In tracking the progress of obtaining more detailed "Program A" cost reports for OSD, the EEVC measure was applied to several program cost reports and outlines:

- "Program A" original (Milestone II) cost estimate
- A cost report on "Program B"
- A cost estimate on Program "C"
- The MIL-STD-881 appendix applicable to "Program A" (outline)
- The first OSD request for additional "Program A" data (outline)
- "Program A" response to the first request
- The second OSD request for additional "Program A" data (outline)
- "Program A" response to the second request

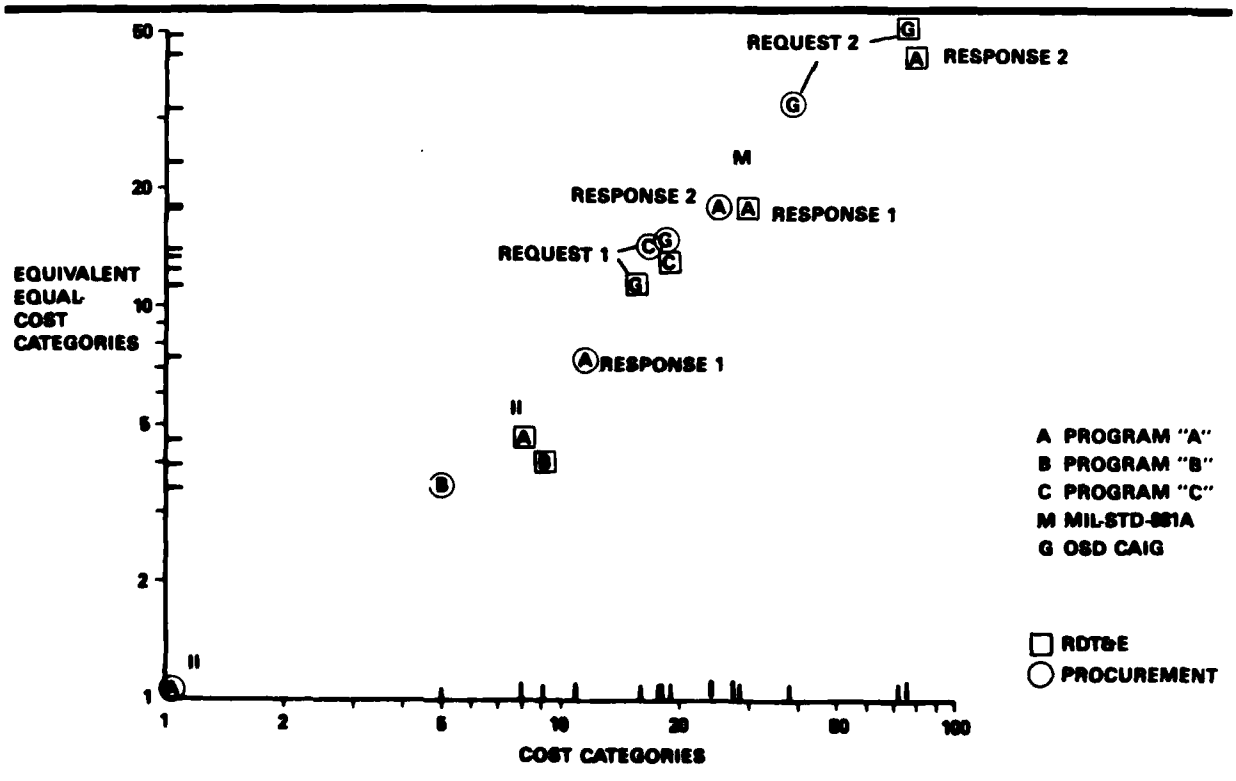
All of the cost reports and outlines, except MIL-STD-881, were examined in two parts: one for RDT&E costs, one for procurement costs. The following table summarizes the detail measured in the subject cost reports and outlines:

INFORMATION MEASURE APPLIED TO PROGRAM "A"

COST REPORT OR OUTLINE	RDT&E #CATEGORIES/#EEVC		PROCUREMENT #CATEGORIES/#EEVC
PROGRAM "A" MILESTONE II POE	8/ 4.5		1/ 1.0
PROGRAM "B" ACTUAL COSTS	9/ 4.0		5/ 3.5
PROGRAM "C" COST ESTIMATE	19/12.7		18/13.6
MIL-STD-881A	—/—	28/23.2	—/—
FIRST OSD CAIG DATA REQUEST	16/11.3		18/14.0
PROGRAM "A" SPO RESPONSE	28/17.5		11/ 7.3
SUBSEQUENT OSD CAIG DATA REQUEST	72/48.9		38/31.7
LATEST PROGRAM "A" SPO RESPONSE	77/42.5		25/17.6

And the following is a graphical display of the above information:

INFORMATION MEASURE APPLIED TO PROGRAM "A"



These diagrams make it easy to notice several things about "Program A" cost reports. The amount of detail in them has increased greatly since the

original Milestone II cost report. Whereas the Milestone II report was generally less informative than cost reports on programs "B" or "C," the latest "Program A" report has more detail than them. Likewise, the (Level 3) cost outline in MIL-STD-881 is approximately as detailed as the latest "Program A" cost report -- less detailed than the RDT&E portion, but more so than the procurement portion. The Service response to the RDT&E portion of the OSD requests has been "better" than for procurement: the first RDT&E response was actually more detailed than the OSD request, while each procurement response has had about a factor-of-two less detail than the OSD request. RDT&E reports and outlines tend to be farther from the equal-sized category "ideal" than are their corresponding procurement reports and outlines. Finally, substantial inequality between cost report categories persists in this data set, regardless of the number of categories. (In the second "Program A" procurement response, there is one large cost category, called "Other," containing 20% of the total procurement cost.)

Limitations, Implementation,and Further Applications of the EEVC Measure.

Limitations of the EEVC measure are easy to state, and generally derive from its abstract nature. One problem is that absolute values — numbers of dollars -- are not reflected. EEVC simply works on the percentage sizes of various cost categories. If any dollar weighting needs to be applied — such as "categories of $100 million or larger" — it must be done outside the EEVC measure. EEVC does not reflect any meaning or usefulness of cost categories. EEVC might "prefer" a five-category report divided into costs of an airplane's red, blue, green, black, and unpainted parts, to a three-category cost report of airframe, avionics, and engines! It remains the cost analyst's responsibility to define cost categories that are meaningful, as well as being numerous and disaggregated.

The EEVC measure is implemented (September, 1985) as a 200-line PASCAL program for a Z-150 microcomputer (an IBM-PC compatible). The user creates an ASCII text file of the input data, according to the following format:

```
Line 1:  Descriptive title of cost outline (65 characters).
Line 2:  1st entry of cost outline (real value, 30-character name),
.
.
.
Line N+1: Last (Nth) entry of cost outline.
End-of-File
```

The user can then run the cost measure program, "EEVC," which asks for the name of the ASCII text file, and whether monitor and/or printer output is desired. EEVC reads the text file and converts it to a data file on disk. EEVC then calculates the degree of detail in the cost outline, in binary bits and in units of equivalent equal-value categories, and lists the input data and the calculated results on the output device.

There are several possibilities for extending the EEVC measure beyond total-program cost reports. Contractor cost report data (e.g., CPR, CCDR) may be examined. Multi-program cost reports (e.g., the FYDP data base) may be studied for inhomogenities or for trends. One challenging area is interpreting time-phased data in the EEVC sense.

END

FILMED

12-85

DTIC